Genomics 98 (2011) 412–421

Contents lists available at ScienceDirect

Genomics

journal homepage: www.elsevier.com/locate/ygeno

Almost total protection from age-related macular degeneration by haplotypes of the Regulators of Complement Activation

Joseph F. Williamson [a,b], Craig A. McLure [a,b,c,d], Robyn H. Guymer [e], Paul N. Baird [e], John Millman [a], Stuart Cantsilieris [c], Roger L. Dawkins [a,b,f,*]

[a] C.Y. O'Connor ERADE Village Foundation, Canning Vale, Western Australia, Australia
[b] Division of Health Sciences, Murdoch University, Murdoch, Western Australia, Australia
[c] Genetic Technologies Limited, Fitzroy, Victoria, Australia
[d] Department of Veterinary Sciences, University of Melbourne, Victoria 3000, Australia
[e] Centre for Eye Research Australia, University of Melbourne, Victoria, Australia
[f] Faculty of Medicine & Dentistry, University of Western Australia, Nedlands, Western Australia, Australia

ARTICLE INFO

Article history:
Received 23 May 2011
Accepted 1 August 2011
Available online 9 August 2011

Keywords:
Age-related macular degeneration
Genetic predisposition to disease
Haplotypes
Ancestral haplotypes
Regulators of Complement Activation
Complement factor H

ABSTRACT

Age-related macular degeneration (AMD) is the leading cause of blindness in developed countries. It has been proposed that the polymorphism encoding Y402H (T1277C) in the complement factor H gene (CFH) is one of the main determinants of disease. We genotyped the polymorphism at a number of loci in the region encompassing the Regulators of Complement Activation (RCA) on chromosome 1, including T1277C SNP, in 187 patients and 146 controls. Haplotypes have been classified as protective (P) or susceptible (S) with respect to AMD. This included the identification of an S haplotype with a T at 1277. The results show that no single locus should be assumed to be directly responsible for AMD, but rather argue for the existence of RCA haplotypes, which can be assigned meaningful predictive values for AMD. We conclude that the critical sequences are within a region 450 kb centromeric to 128 kb telomeric of CFH.

1. Introduction

Polygenic diseases, such as Age-Related Macular Degeneration (AMD), have proven difficult to understand, at least partly because of the complexity of the interactions between linked and unlinked genes, the influence of noncoding polymorphisms affecting gene regulation, variable age and sex related penetrance, and diverse environmental factors [1–3]. These problems are compounded when the pathogenesis is poorly understood.

AMD is one polygenic disease which appears to be amenable to progress. There are at least several gene clusters involved, but interestingly most (complement factor H (CFH), complement component 2 (C2), complement factor B (CFB), complement component 3 (C3), mannose-binding lectin 2 (MBL2)) may relate to complement activation within vessels and may thus be understood in terms of functional interactions and quantitative effects [4–12]. At least some risk factors are well known and contribute to a final common pathway involving inflammation and obstruction of terminal vasculature. Effective therapy with anti vascular endothelial growth factor (VEGF) is available [13] although genetic assays to identify those requiring treatment would be helpful both clinically and financially.

An unresolved series of problems relate to the multiplicity and complexity of the genetic markers in the vicinity of CFH and to the presence of some which may or may appear to contribute to either susceptibility or protection. As one example of the confusion, Spencer et al. [5] report contradictory results and suggest the use of the term "inverse associations" until the pathophysiology is known. The same group also discussed the difficulties resulting from multiple comparisons and from inferring rather than observing haplotypes [14].

A proven approach to evaluating disease associations has been developed through studies of the Major Histocompatibility Complex (MHC) [15–21]. The identification of ancestral haplotypes (AH) provides a means of studying the extensive coding and noncoding sequences which have been selected through human evolution. These are the sequences which are likely to be most important in function and disease. Ancestral haplotypes contain genomic structure, insertions and deletions (indels), copy number variations (CNVs), and regulatory elements as well as single nucleotide polymorphisms (SNPs) and microsatellites. Thus potential complexity can be simplified resulting in

Abbreviations: AH, ancestral haplotype; AMD, age-related macular degeneration; C2, complement component 2; C3, complement component 3; CFB, complement factor B; CFH, complement factor H; CNV, copy number variation; indels, insertions or deletions; MBL2, mannose-binding lectin 2; MHC, major histocompatibilty complex; P, protective haplotype; RCA, Regulators of Complement Activation; S, susceptible haplotype; VEGF, vascular endothelial growth factor.

* Corresponding author at: CY O'Connor ERADE Village, PO Box 5100, Canning Vale South, Western Australia 6155, Australia. Fax: +61 8 9397 1559.
E-mail address: rldawkins@cyo.edu.au (R.L. Dawkins).

0888-7543/$ – see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.ygeno.2011.08.002

Table 1
Frequency and size of FH1 alleles in 187 AMD patients and 146 controls.

Size (bp)	FH1 allele	Allele frequency (%)
530	3	0·3
533	4	0·5
535	5	0.2
537	6	1·5
541	7	1·8
543	8	3·6
545	9	5·4
547	10	5·56
549	11	10·1
551	12	5·56
553	13	6·0
555	14	6·3
557	15	4·8
559	16	12·9
562	17	3·9
564	18	10·2
566	19	2·1
568	20	5·56
570	21	0·9
572	22	1·0
574	23	0·2
576	24	0·6
580	25	0·3
600	26	0·2
605	27	1·0
609	28	1·95
613	29	1·8
617	30	1·5
621	31	2·1
625	32	1·8
628	33	0·4

a powerful strategy for examining genetic interactions in determining susceptibility to disease.

Recently [22], we have characterised haplotypes of the beta block of the RCA [23]. Here, we show that there are many more haplotypes than described to date and that these can be classified as protective (P) or susceptible (S) with respect to AMD. The P and S haplotypes differ in several respects suggesting some explanations for the disease association.

A potential clinical application is suggested by the fact that the PP genotype is present in only 2% patients with AMD but 24% of controls. Thus, a PP genotype may be useful in excluding the need for intensive follow-up and treatment.

2. Results and discussion

2.1. The polymorphic FH1 locus as the basis for classifying ancestral haplotypes

The FH1 locus is highly polymorphic, with more than 30 alleles found in the 333 subjects typed to date. The more frequent alleles, designated 7 through to 24, differ in amplicon length by approximately 2 base pairs such that allele 7 is 541 bp, 8 is 543 bp, 9 is 545 bp, 10 is 547 bp etc., as shown in Table 1.

To facilitate comparisons, the FH1 genotypes of all patients and controls are displayed in a percentile database (Fig. 1). Subjects are listed in approximate order of the gene frequency when the two groups are combined. Patients with AMD are listed on the left together with their clinical type and age. Controls are on the right. Most subjects have 2 different alleles; those with only one are assumed to be homozygous.

It is immediately clear that the AMD and controls have very different allele frequencies. For example, in AMD, there are increases in FH1 16, 18, 14, 10, 20 and 12 but decreases in 11, 13, 9 and 17. Note that the alleles increased in AMD are a set (10, 12, 14, 16, 18) reflecting a 4 base difference in length (see Table 1). By contrast, the decreased alleles belong to a different set (9, 11, 13, 17) also reflecting a 4 base difference in length, but offset by 2 bases as shown in Table 1. The exception, FH1 15, is considered further below.

The alternating pattern of frequency is illustrated in Fig. 2 where it can be seen that most alleles can be classified into one of the 2 groups provisionally designated susceptibility (S) or protective (P) alleles. Accordingly, each genotype could be assigned SS, SP or PP and in some cases SX, PX or XX where the effect of X on AMD susceptibility is unknown due to low allele frequency.

To test the validity of this approach we examined the *trans*-interactions between S and P alleles. Initially we excluded lower frequency alleles and selected the patients and controls with genotypes restricted to the P alleles (11, 13) and S alleles (12 ,16, 20) which are most significantly increased or decreased as shown in Table 2 and Fig. 3. Seventeen of the 21 controls have PP or PS. None of the patients has PP. Note the PP, PS, SS distributions of 0·29, 0·52, 0·19 in controls and 0, 0·37, 0·63 in AMD. These results suggest that P alleles are indeed protective when homozygous and that S alleles are at least additive.

Very similar results were obtained when the 12 most frequent alleles were included in the analysis (Fig. 4). The P alleles (9, 11, 13, 17) and S alleles (8, 10, 12, 14, 15, 16, 18, 20) gave PP, PS, SS distributions of 0·28, 0·47, 0·24 in controls and 0·02, 0·37, 0·61 in AMD. Only 3 of 137 patients are homozygous PP.

The classification of lower frequency alleles given in Table 1 and Fig. 2 is provisional but, when based on the odd and even rule, the results are again similar with a maximal estimate for PP of 0·06 in AMD. In controls, the gene frequencies of P and S again approximate 0·5 in Hardy Weinberg Equilibrium.

2.2. Other loci define S and P haplotypes

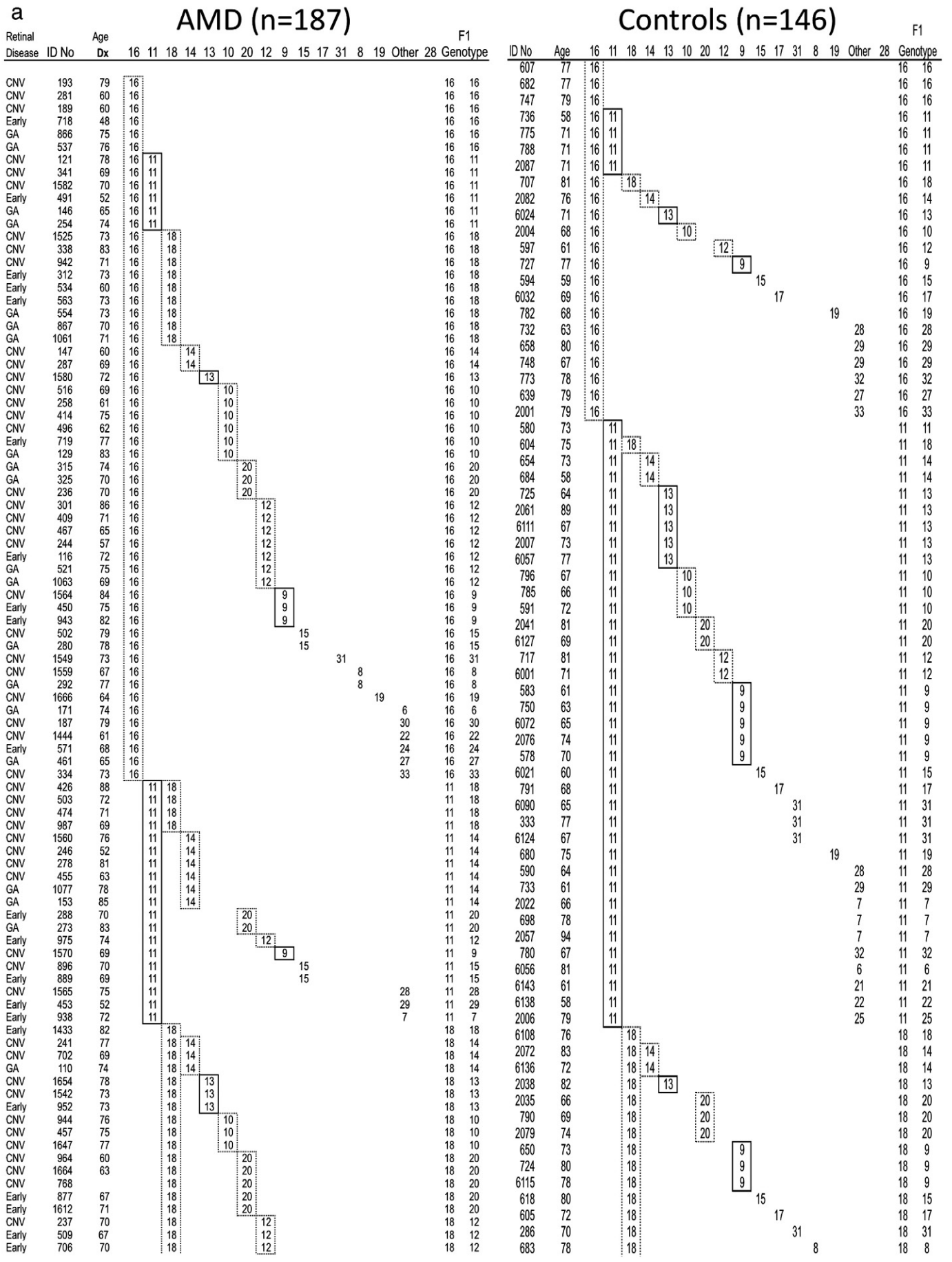
Since it is known that the T1277C mutation is associated with AMD and located close to the FH1 locus, we examined the relationship between FH1 alleles and 1277 alleles. All of the S alleles (such as 10, 12, 14, 16, 18) carry C, whereas P alleles (9, 11, 13, 17) carry T. Interestingly, FH1 15 also has a T although it is increased rather than decreased in AMD and therefore cannot be classified as a P allele. In fact, 8 of the 18 AMD patients with the TT genotype have FH1 15, confirming that this haplotype is distinctive. This finding suggests that neither the FH1 nor the 1277 locus is directly involved in susceptibility or protection: alleles may merely be imperfect markers in linkage disequilibrium with other components on critical haplotypes.

In Fig. 5, the haplotypes including FH1 15 and T1277C are shown. Note that 8 of the patients have TT confirming that TT is not protective in the presence of FH1 15. There are numerous potential haplotypes when the −45 and +128 loci are included. Note the frequencies of −45 6 (8/19 v 2/11) and +128 2 (11/19 v 3/11) alleles are the principal differences between AMD and controls suggesting that the −45 and +128 loci may be as important as the FH1 or 1277 loci in marking the critical factors involved in protection and susceptibility to AMD. If so, it is important to define the haplotypes rather than alleles at any single locus.

In a previous study based on 3 generation families [22], 10 ancestral haplotypes of −45, 1277, FH1 and +128 were identified including 5, T, 15, 1. As shown in Fig. 5, the patients have 6, T, 15, 1 or 5, T, 15, 2 rather than 5, T, 15, 1, suggesting that recombination

Fig. 1. Genotypes of all individuals in the study represented as a percentile database. Alleles are listed on the top row. Subjects are listed in approximate order of the gene frequency (when the 2 groups are combined). Patients with AMD are listed on the left together with the clinical type and age. Controls are shown on the right. Most subjects have 2 different alleles; those with only one are assumed to be homozygous. Alleles with a frequency>5% (Table 2) in the combined groups have been boxed. The alleles increased in patients are boxed with a dashed line. Alleles decreased in patients are boxed with a solid line.

a


AMD (n=187)
Controls (n=146)
Retinal Disease
ID No
Age Dx
F1 Genotype
CNV 193 79 16 16 16
CNV 281 60 16 16 16
CNV 189 60 16 16 16
Early 718 48 16 16 16
GA 866 75 16 16 16
GA 537 76 16 16 16
CNV 121 78 16 11 16 11
CNV 341 69 16 11 16 11
CNV 1582 70 16 11 16 11
Early 491 52 16 11 16 11
GA 146 65 16 11 16 11
GA 254 74 16 11 16 11
CNV 1525 73 16 18 16 18
CNV 338 83 16 18 16 18
CNV 942 71 16 18 16 18
Early 312 73 16 18 16 18
Early 534 60 16 18 16 18
Early 563 73 16 18 16 18
GA 554 73 16 18 16 18
GA 867 70 16 18 16 18
GA 1061 71 16 18 16 18
CNV 147 60 16 14 16 14
CNV 287 69 16 14 16 14
CNV 1580 72 16 13 16 13
CNV 516 69 16 10 16 10
CNV 258 61 16 10 16 10
CNV 414 75 16 10 16 10
CNV 496 62 16 10 16 10
Early 719 77 16 10 16 10
GA 129 83 16 10 16 10
GA 315 74 16 20 16 20
GA 325 70 16 20 16 20
CNV 236 70 16 20 16 20
CNV 301 86 16 12 16 12
CNV 409 71 16 12 16 12
CNV 467 65 16 12 16 12
CNV 244 57 16 12 16 12
Early 116 72 16 12 16 12
GA 521 75 16 12 16 12
GA 1063 69 16 12 16 12
CNV 1564 84 16 9 16 9
Early 450 75 16 9 16 9
Early 943 82 16 9 16 9
CNV 502 79 16 15 16 15
GA 280 78 16 15 16 15
CNV 1549 73 16 31 16 31
CNV 1559 67 16 8 16 8
GA 292 77 16 8 16 8
CNV 1666 64 16 19 16 19
GA 171 74 16 6 16 6
CNV 187 79 16 30 16 30
CNV 1444 61 16 22 16 22
Early 571 68 16 24 16 24
GA 461 65 16 27 16 27
CNV 334 73 16 33 16 33
CNV 426 88 11 18 11 18
CNV 503 72 11 18 11 18
CNV 474 71 11 18 11 18
CNV 987 69 11 18 11 18
CNV 1560 76 11 14 11 14
CNV 246 52 11 14 11 14
CNV 278 81 11 14 11 14
CNV 455 63 11 14 11 14
GA 1077 78 11 14 11 14
GA 153 85 11 14 11 14
Early 288 70 11 20 11 20
GA 273 83 11 20 11 20
Early 975 74 11 12 11 12
CNV 1570 69 11 9 11 9
CNV 896 70 11 15 11 15
Early 889 69 11 15 11 15
CNV 1565 75 11 28 11 28
Early 453 52 11 29 11 29
Early 938 72 11 7 11 7
Early 1433 82 18 18 18 18
CNV 241 77 18 14 18 14
CNV 702 69 18 14 18 14
GA 110 74 18 14 18 14
CNV 1654 78 18 13 18 13
CNV 1542 73 18 13 18 13
Early 952 73 18 13 18 13
CNV 944 76 18 10 18 10
CNV 457 75 18 10 18 10
CNV 1647 77 18 10 18 10
CNV 964 60 18 20 18 20
CNV 1664 63 18 20 18 20
CNV 768 18 20 18 20
Early 877 67 18 20 18 20
Early 1612 71 18 20 18 20
CNV 237 70 18 12 18 12
Early 509 67 18 12 18 12
Early 706 70 18 12 18 12
ID No Age
F1 Genotype
607 77 16 16 16
682 77 16 16 16
747 79 16 16 16
736 58 16 11 16 11
775 71 16 11 16 11
788 71 16 11 16 11
2087 71 16 11 16 11
707 81 16 18 16 18
2082 76 16 14 16 14
6024 71 16 13 16 13
2004 68 16 10 16 10
597 61 16 12 16 12
727 77 16 9 16 9
594 59 16 15 16 15
6032 69 16 17 16 17
782 68 16 19 16 19
732 63 16 28 16 28
658 80 16 29 16 29
748 67 16 29 16 29
773 78 16 32 16 32
639 79 16 27 16 27
2001 79 16 33 16 33
580 73 11 11 11 11
604 75 11 18 11 18
654 73 11 14 11 14
684 58 11 14 11 14
725 64 11 13 11 13
2061 89 11 13 11 13
6111 67 11 13 11 13
2007 73 11 13 11 13
6057 77 11 13 11 13
796 67 11 10 11 10
785 66 11 10 11 10
591 72 11 10 11 10
2041 81 11 20 11 20
6127 69 11 20 11 20
717 81 11 12 11 12
6001 71 11 12 11 12
583 61 11 9 11 9
750 63 11 9 11 9
6072 65 11 9 11 9
2076 74 11 9 11 9
578 70 11 9 11 9
6021 60 11 15 11 15
791 68 11 17 11 17
6090 65 11 31 11 31
333 77 11 31 11 31
6124 67 11 31 11 31
680 75 11 19 11 19
590 64 11 28 11 28
733 61 11 29 11 29
2022 66 11 7 11 7
698 78 11 7 11 7
2057 94 11 7 11 7
780 67 11 32 11 32
6056 81 11 6 11 6
6143 61 11 21 11 21
6138 58 11 22 11 22
2006 79 11 25 11 25
6108 76 18 18 18 18
2072 83 18 14 18 14
6136 72 18 14 18 14
2038 82 18 13 18 13
2035 66 18 20 18 20
790 69 18 20 18 20
2079 74 18 20 18 20
650 73 18 9 18 9
724 80 18 9 18 9
6115 78 18 9 18 9
618 80 18 15 18 15
605 72 18 17 18 17
286 70 18 31 18 31
683 78 18 8 18 8

b

AMD (n=187)

Retinal Disease	ID No	Age Dx	16	11	18	14	13	10	20	12	9	15	17	31	8	19	Other	28	F1 Genotype	
CNV	173	79			18							15							18	15
CNV	339	80			18							15							18	15
Early	493	62			18							15							18	15
Early	835	68			18								17						18	17
Early	454	68			18								17						18	17
Early	101	73			18								17						18	17
Early	951	85			18										8				18	8
Early	1547	65			18										8				18	8
GA	894	62			18										8				18	8
CNV	337	83			18												28		18	28
GA	488	66			18												24		18	24
CNV	1548	81			18												27		18	27
CNV	249	72				14													14	14
CNV	1540	90				14		10											14	10
CNV	1443	68				14		10											14	10
Early	1080	78				14				12									14	12
Early	1088	72				14				12									14	12
Early	905	61				14				12									14	12
CNV	845	81				14					9								14	9
Early	1650	79				14					9								14	9
Early	673	64				14						15							14	15
GA	1436	70				14							17						14	17
CNV	476	77				14									8				14	8
CNV	1556	63				14									8				14	8
CNV	1546	72				14										19			14	19
Early	492	80				14											29		14	29
CNV	1519	72				14											32		14	32
CNV	1649	57				14													14	21
CNV	1044	83					13										21		13	13
CNV	470	72					13												13	13
GA	843	70					13		20										13	20
CNV	1557	78					13					15							13	15
CNV	857	63					13							31					13	31
Early	1579	73					13												13	22
CNV	284	69					13										22		13	24
CNV	247	77						10									24		10	10
CNV	264	62						10											10	10
CNV	848	70						10		12									10	12
CNV	1607	84						10			9								10	9
CNV	1555	65						10			9								10	9
CNV	1059	67						10				15							10	15
CNV	766	72						10					17						10	17
CNV	1563	75						10							8				10	8
Early	1402	70						10							8				10	8
CNV	239	57						10									7		10	7
CNV	125	73						10									32		10	32
CNV	114	67						10											10	27
CNV	106	76							20	12							27		20	12
CNV	155	71							20	12									20	12
Early	452	73							20	12									20	12
GA	1543	85							20		9								20	9
CNV	743	75							20			15							20	15
Early	1429	76							20			15							20	15
Early	851	73							20				17						20	17
CNV	1522	71							20						8				20	8
Early	501	83							20						8				20	8
CNV	499	74							20							19			20	19
CNV	196	75							20							19			20	19
CNV	327	85							20								28		20	28
CNV	945	71							20								28		20	28
CNV	1439	75							20										20	3
CNV	1414	89							20								3		20	21
Early	456	79							20								21		20	22
CNV	107	68								12							22		12	12
Early	1576	78								12	9								12	9
Early	439	63								12	9								12	9
Early	517	65								12	9								12	9
CNV	752	66								12			17						12	17
CNV	1655	77								12				31					12	31
CNV	170	73								12							32		12	32
CNV	182	80								12							6		12	6
GA	672	68								12							30		12	30
CNV	389	82								12									12	3
CNV	285	83								12							3		12	22
CNV	1550	68									9	15					22		9	15
CNV	767	70									9	15							9	15
CNV	282	64									9				8				9	8
CNV	976	71									9					19			9	19
GA	955	70									9						28		9	28
CNV	947	79										15							15	15
GA	1541	70										15							15	15
CNV	1523	77										15	17						15	17
Early	500	70										15			8				15	8
CNV	868	77										15							15	22
CNV	152	57											17		8		22		17	8
CNV	1538	86											17				32		17	32
Early	1089	67											17				30		17	30
CNV	122	70											17				21		17	21
Early	302	69												31			21		31	21
CNV	936	72													8		28		8	28
Early	176	61													8		4		8	4
Early	346	76													8		25		8	25
GA	331	78													8		22		8	22
CNV	755	75															7	32	7	32
Early	425	65															23	24	23	24

Controls (n=146)

ID No	Age	16	11	18	14	13	10	20	12	9	15	17	31	8	19	Other	28	F1 Genotype	
6110	75			18											19			18	19
6095	65			18											19			18	19
2016	85			18												7		18	7
6015	66			18												7		18	7
2030	75			18												6		18	6
2008	87			18												27		18	27
6016	76			18												27		18	27
730	77				14	13												14	13
681	78				14					9								14	9
627	79				14							17						14	17
6128	76				14									8				14	8
593	78				14											28		14	28
608	72				14											29		14	29
6144	64				14											7		14	7
595	64				14											30		14	30
2044	71				14											33		14	33
6067	85					13												13	13
783	72					13	10											13	10
2049	82					13	10											13	10
795	67					13	10											13	10
2013	80					13		20										13	20
787	64					13			12									13	12
6107	80					13				9								13	9
677	75					13					15							13	15
6088	66					13						17						13	17
697	68					13						17						13	17
6130	65					13							31					13	31
644	77					13							31					13	31
6069	69					13										28		13	28
2063	86					13										7		13	7
651	68					13										32		13	32
2043	77					13										6		13	6
2089	77					13										6		13	6
6129	65					13										30		13	30
708	70						10			9								10	9
632	75						10				15							10	15
621	66						10									32		10	32
2060	75						10									4		10	4
589	73						10									27		10	27
6009	61							20			15							20	15
586	66							20			15							20	15
6109	87							20				17						20	17
716	72							20								29		20	29
777	66								12							29		12	29
6122	67								12							6		12	6
6073	80								12							30		12	30
6071	58									9								9	9
592	58									9		17						9	17
2042	80									9		17						9	17
6030	80									9			31					9	31
2088	77									9				8				9	8
6137	61									9						28		9	28
2074	63										15				19			15	19
2058	77										15					7		15	7
617	70										15					32		15	32
712	69										15					26		15	26
711	78											17	31					17	31
6034	72											17			19			17	19
784	61											17				28		17	28
6142	66											17				6		17	6
2033												17				4		17	4
723	72												31			29		31	29
2081	74												31			30		31	30
2026														8		28		8	28
587	58													8		29		8	29
713	60													8		30		8	30
786	71														19	29		19	29
638	76														19	7		19	7
2011	84														19	6		19	6
2005	75															30	29	29	30
2014	72															30	32	32	30
2077	80															4	32	32	4
600	71															5	6	6	5

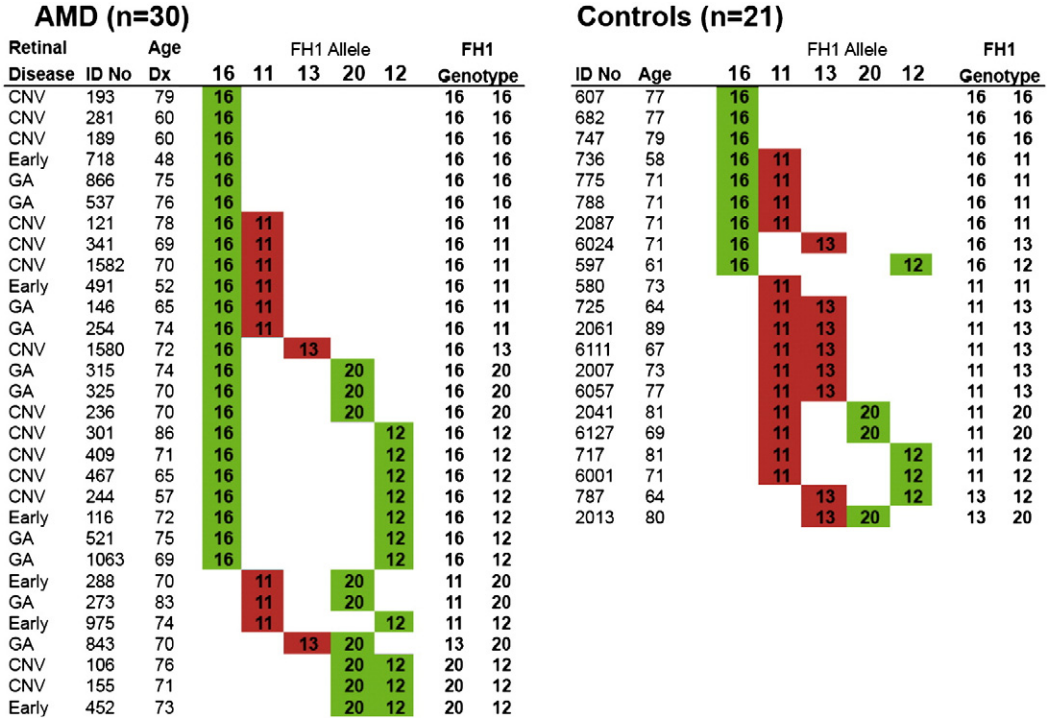


Fig. 2. The allele frequencies (%) of the common FH1 alleles in AMD patients (n = 187) and controls (n = 146). Alleles designated susceptible are shown in green and protective alleles are shown in red. The unbroken line represents the combined (patients and controls) allele frequency.

centromeric or telomeric of CFH is responsible for the increased frequencies of −45 6 and +128 2 in disease and for susceptibility rather than protection. Note that −45 6, in the absence of T, 15, is not associated with disease, therefore implying that there are *cis*-interactions with unidentified sequences marked by −45 6.

2.3. AMD and diabetes

In a subset of the patients with AMD (n = 46), the presence or absence of type 2 diabetes (T2D) was known. The frequencies of FH1 in these patients are given in Fig. 6 which suggests that FH1 P alleles such as 9 and 11 may not be protective for AMD in the presence of diabetes. As expected, the frequencies in AMD patients without diabetes are very similar to the total AMD group.

Table 2
The F1 allele defines haplotypes with different levels of protection/susceptibility for AMD. Only F1 alleles with frequencies >5% in the combined groups are shown.

F1 allele	Controls 2n† = 292	AMD 2n† = 374	Combined groups freq (%)	Odds ratio	p‡
12	7	30	5·56	3·55	0·002
20	10	27	5·56	2·19	0·04
16	25	61	12·91	2·08	0·003
10	12	25	5·56	1·67	0·174
14	14	28	6·31	1·61	0·199
18	24	44	10·21	1·49	0·156
9	19	17	5·41	0·68	0·302
11	42	25	10·06	0·43	0·001
13	27	13	6·01	0·35	0·003

2n† = Number of haplotypes.
‡Fisher's exact 2-tailed probability.

2.4. Clinical utility of FH1 and T1277C in excluding AMD

There have been suggestions that homozygous TT at 1277 may be helpful in excluding AMD. However, in our current data set, the negative predictive value of TT is too low to permit critical clinical decisions. The benefit of PP over TT as an excluding test is largely due to the fact that the 6 TT genotypes including FH1 15 become PS (see Fig. 4).

The present results argue for the existence of RCA beta haplotypes which carry the sequences responsible for protection or susceptibility to AMD. No single locus, including 1277, should be assumed to be directly responsible.

It is always hazardous to assign protection and susceptibility alleles to a single locus since increased frequency of some alleles must be compensated by reductions in others. In the present study, however, we are confident that there are fundamental differences between S and P haplotypes because, with the exception of FH1 15:

1. The distribution of SS, SP and PP in disease is not a simple function of frequency.
2. *Trans*-interactions between S and P haplotypes are clearly powerful, with S alleles dominating over P, although undoubtedly complex.

In general, we are reluctant to ascribe significance to HW distortion in disease since this often reflects an inability to detect heterozygotes. This is not the case in the present study.

Furthermore, with the exception of FH 15:

1. S and P alleles relate to sets determined by amplicon length.
2. S and P alleles relate to T1277C

3. Conclusions

Accordingly, we conclude that there are critical sequences in the vicinity of the 1.7 Mb region between −45 to +128 and including

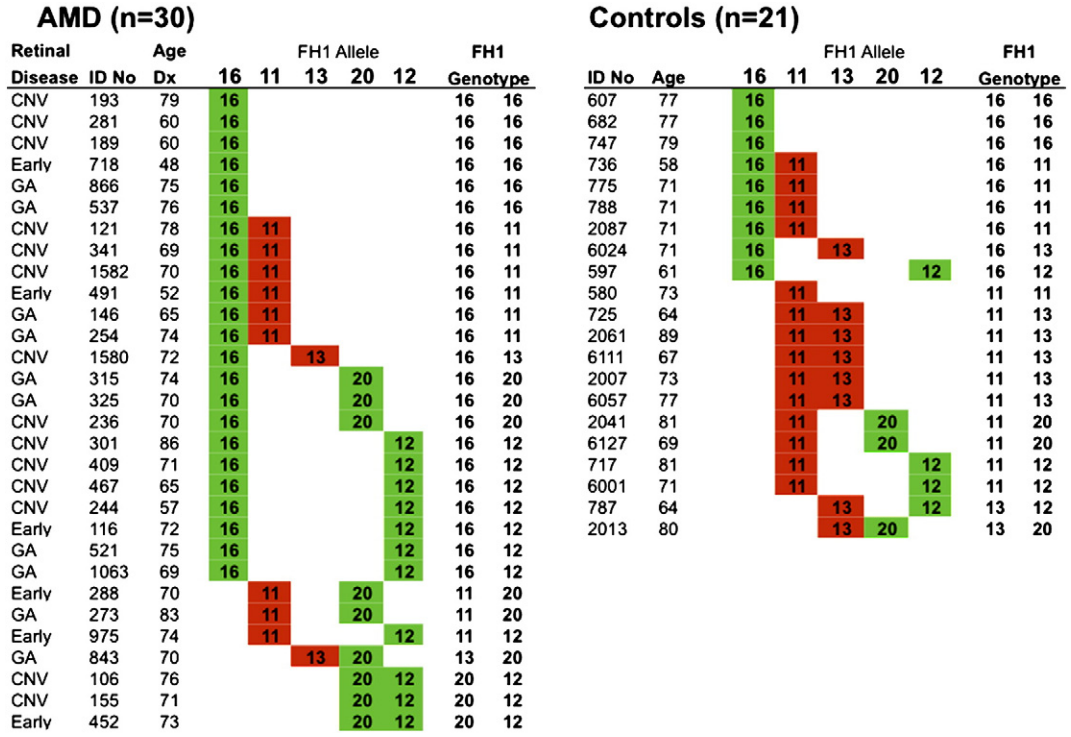


Fig. 3. FH1 genotype distribution in a subset of AMD and control individuals that carry combinations of common alleles that show statistically significant, positive or negative, association with AMD. Susceptible alleles are shown in green. Protective alleles are in red.

CFH and we postulate that there are multiple *cis*-interactions to explain the current results.

These sequences profoundly influence susceptibility to AMD. Subjects with PP genotypes such as 11, 13 have a very low risk of disease and sufficiently low to suggest that genetic typing may be useful clinically. For the first time it becomes possible to assign meaningful predictive values or risks to particular genotypes. We anticipate that further characterisation of RCA beta haplotypes will enhance clinical utility.

Previously, we compared the sequences amplified from FH1 16 and 17 homozygotes and found that 16 and other S haplotypes had a deletion of 2 bp plus multiple substitutions representing an SNP rate of some 5% [22]. The 2 bp indel is found in sequences with a T at 1277 and appears to have arisen contemporaneously during primate evolution. The length differences are largely a function of the number of copies of diverse C and T rich tetramers inserted and deleted on either side of the 2 bp indel. There are also more extensive indels associated with S and P haplotypes.

Future studies will quantify the role and interactions between haplotypes, examine the effect of more extensive RCA beta haplotypes on therapeutic responses and perhaps assist in determining which cases require early treatment.

In parallel with improved genetic typing, it will be valuable to include other risk factors which may overcome the protective effect of the PP genotype. A possible example is provided in the present data; diabetes may predispose to AMD in PP subjects, but further studies are required to confirm these results and determine whether the effect is due to diabetes *per se* or further genetic interactions.

As argued here, the mechanism of disease protection is probably quantitative and regulatory, since certain haplotypes are associated with greater or lesser degrees of inflammation and vascular injury. However qualitative changes in CFH itself may be also be important. It seems likely that other products encoded in the region and in other regions, such as the MHC, may also contribute. In these regards, AMD will serve as a useful model for dissecting the contributions to polygenic diseases.

4. Materials and methods

4.1. Strategy

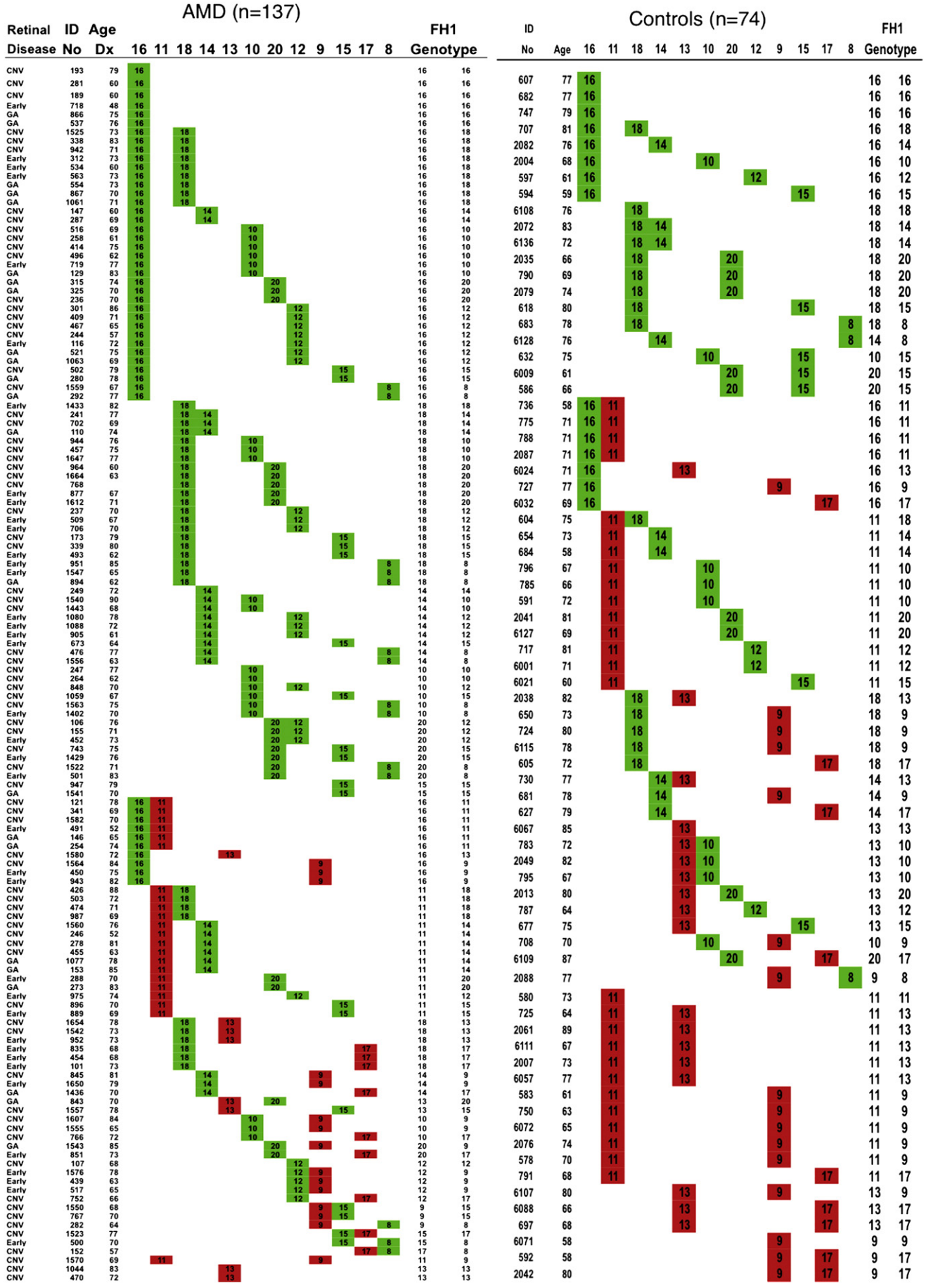
The intention is to determine the coding and noncoding sequences which together determine susceptibility to, or protection from, a polygenic disease. This resolves into the following steps.

Firstly, it is necessary to define multiple linked polymorphisms, and particularly those which have numerous alleles, which can be detected reliably. Combinations of alleles at contiguous loci allow the recognition of candidate haplotypes. To have any value in studying disease and control populations, these combinations must be observed rather than inferred [24] and must be maintained through successive matings. Operationally, an ancestral haplotype is defined as a combination of alleles which occurs:

1. as half the content of at least 3 different genotypes defined in a family study including at least 3 generations,
2. in at least 3 subjects not likely to be related by recent descent.

In such cases, it has been found that the entire sequence between and beyond the component loci is identical, with the important exception that some rapidly mutating microsatellites will exhibit slippage. Thus the haplotype marks indels, CNVs, stable microsatellites and SNPs and can be numbered generally by reference to the component allele at the most polymorphic locus, with a second digit to reflect the order of discovery e.g. MHC AH8.1 [25].

Secondly, the extent of the haplotype is determined by locating the boundaries defined by recombination. As multiple examples are accumulated, it is often possible to identify rare recombinants between


AMD (n=137)
Retinal Disease
ID No
Age Dx
16 11 18 14 13 10 20 12 9 15 17 8
FH1 Genotype
Controls (n=74)
ID No
Age
16 11 18 14 13 10 20 12 9 15 17 8
FH1 Genotype

Retinal Disease	AMD ID No	-45		1277		FH1		+128	
Early	1429	3	7	T	C	15	20	2	1
CNV	502	3	2	T	C	15	16	2	1
CNV	339	3	3	T	C	15	18		
CNV	1523	3	1	T	T	15	17	1	3
Early	889	4	7	T	T	15	11	1	2
CNV	1557	4	8	T	T	15	13	1	1
CNV	947	5	9	T	T	15	15	2	1
GA	1541	5	4	T	T	15	15	2	1
Early	493	5	8	T	C	15	18	2	2
CNV	743	5	8	T	C	15	20	2	2
Early	500	5	4	T	C	15	8	2	2
CNV	173	6	2	T	C	15	18	1	2
CNV	1059	6	3	T	C	15	10	1	2
CNV	767	6	8	T	T	15	9	1	2
GA	280	6	4	T	C	15	16	1	1
CNV	868	6	8	T	C	15	22	1	1
CNV	1550	6	1	T	T	15	9	1	1
CNV	896	6	6	T	T	15	11	1	1
Early	673	6	1			15	14	1	1

Controls ID No	-45		1277		FH1		+128	
618	3	8	T	C	15	18	1	1
632	3	4	T	C	15	10	1	2
712	3	2	T	T	15	26	1	1
617	3	8	T	T	15	32	1	1
2074	4	5	T	T	15	19	1	1
6009	4	5	T	C	15	20	1	2
586	4	2	T	C	15	20	1	1
6021	4	1	T	T	15	11	1	1
677	5	7	T	T	15	13	2	1
2058	6	3	T	T	15	7		
594	6	1			15	16	1	1

Fig. 5. Alleles at −45 and +128 loci in FH1-15 individuals.

ancestral haplotypes at the more remote sites. These become useful in mapping the contributions of the components [26,27].

Thirdly, it is possible to compare and contrast AHs and to classify groups by virtue of their different effects in terms of susceptibility and protection. *Trans, cis* and epistatic interactions can be defined. The validity of the groupings within the population can be tested using the Hardy Weinberg equation.

In the present study, we have used the AHs defined by Williamson et al. [22] to examine the effects in AMD and particularly the contributions to susceptibility and protection.

4.2. Methods

The genomic region of 1.7 Mb containing CFH and its copies was interrogated with multiple potential markers, including the T1277C mutation in exon 9 of CFH and the FH1 polymorphism in intron 9. The gene content and marker locations are shown in Fig. 7. By also using the −45 and +128 markers ancestral haplotypes could be defined. All methods as Williamson et al. [22].

4.3. Patients and controls

The Centre for Eye Research Australia (RG and PNB) provided 187 AMD patients and 146 aged matched controls, which have been described in previous studies [10,28]. All AMD participants were recruited for the AMD inheritance study (AMDIS) from ophthalmology clinics in Melbourne. Individuals were classified into clinical subsets as 1. early (n = 49), 2. geographic atrophy (GA; n = 29) or subfoveal choroidal neovascularisation (CNV; n = 109). Controls were collected as part of the Visual Impairment Project (VIP), a population-based epidemiological eye study or through aged-care nursing homes. All individuals were of Anglo-Celtic ethnic background. On a subset of 35 control and 46 AMD individuals, additional clinical information was available. This included the presence or absence of type 2 diabetes (T2D).

Informed written consent was obtained from all individuals. Ethics approval was provided by the Human Research and Ethics Committee of the Royal Victorian Eye and Ear Hospital. The study was conducted in accordance with the Declaration at Helsinki.

Acknowledgments

We are grateful to Brent Stewart, Peter Keating and Luisa Ashdown for helpful discussions and Susan Lester for assistance with statistical analyses. This is Manuscript 0702 of the C Y O'Connor ERADE Village Foundation.

References

[1] A.E. Hughes, N. Orr, C. Patterson, H. Esfandiary, R. Hogg, V. McConnell, G. Silvestri, U. Chakravarthy, Neovascular age-related macular degeneration risk based on CFH, LOC387715/HTRA1, and smoking, PLoS Med 4 (2007) e355.
[2] T.E. Clemons, R.C. Milton, R. Klein, J.M. Seddon, F.L. Ferris III, Risk factors for the incidence of Advanced Age-Related Macular Degeneration in the Age-Related Eye Disease Study (AREDS) AREDS report no. 19, Ophthalmology 112 (2005) 533–539.
[3] A. Lotery, D. Trump, Progress in defining the molecular biology of age related macular degeneration, Hum. Genet. 122 (2007) 219–236.
[4] D.A. Schaumberg, W.G. Christen, P. Kozlowski, D.T. Miller, P.M. Ridker, R.Y. Zee, A prospective assessment of the Y402H variant in complement factor H, genetic variants in C-reactive protein, and risk of age-related macular degeneration, Invest. Ophthalmol. Vis. Sci. 47 (2006) 2336–2340.
[5] K.L. Spencer, M.A. Hauser, L.M. Olson, S. Schmidt, W.K. Scott, P. Gallins, A. Agarwal, E.A. Postel, M.A. Pericak-Vance, J.L. Haines, Protective effect of complement factor B and complement component 2 variants in age-related macular degeneration, Hum. Mol. Genet. 16 (2007) 1986–1992.
[6] V. Dinu, P.L. Miller, H. Zhao, Evidence for association between multiple complement pathway genes and AMD, Genet. Epidemiol. 31 (2007) 224–237.
[7] J.L. Haines, M.A. Hauser, S. Schmidt, W.K. Scott, L.M. Olson, P. Gallins, K.L. Spencer, S.Y. Kwan, M. Noureddine, J.R. Gilbert, N. Schnetz-Boutaud, A. Agarwal, E.A. Postel, M.A. Pericak-Vance, Complement factor H variant increases the risk of age-related macular degeneration, Science 308 (2005) 419–421.
[8] K.P. Magnusson, S. Duan, H. Sigurdsson, H. Petursson, Z. Yang, Y. Zhao, P.S. Bernstein, J. Ge, F. Jonasson, E. Stefansson, G. Helgadottir, N.A. Zabriskie, T. Jonsson, A. Bjornsson, T. Thorlacius, P.V. Jonsson, G. Thorleifsson, A. Kong, H. Stefansson, K. Zhang, K. Stefansson, J.R. Gulcher, CFH Y402H confers similar risk of soft drusen and both forms of advanced AMD, PLoS Med. 3 (2006) e5.

Fig. 4. Genotype distribution of FH1 alleles in those patients and controls that carry the 12 most frequent alleles. Susceptible alleles are shown in green. Protective alleles are shown in red. Subjects are sorted by allele frequencies within SS, SP and PP genotypes.

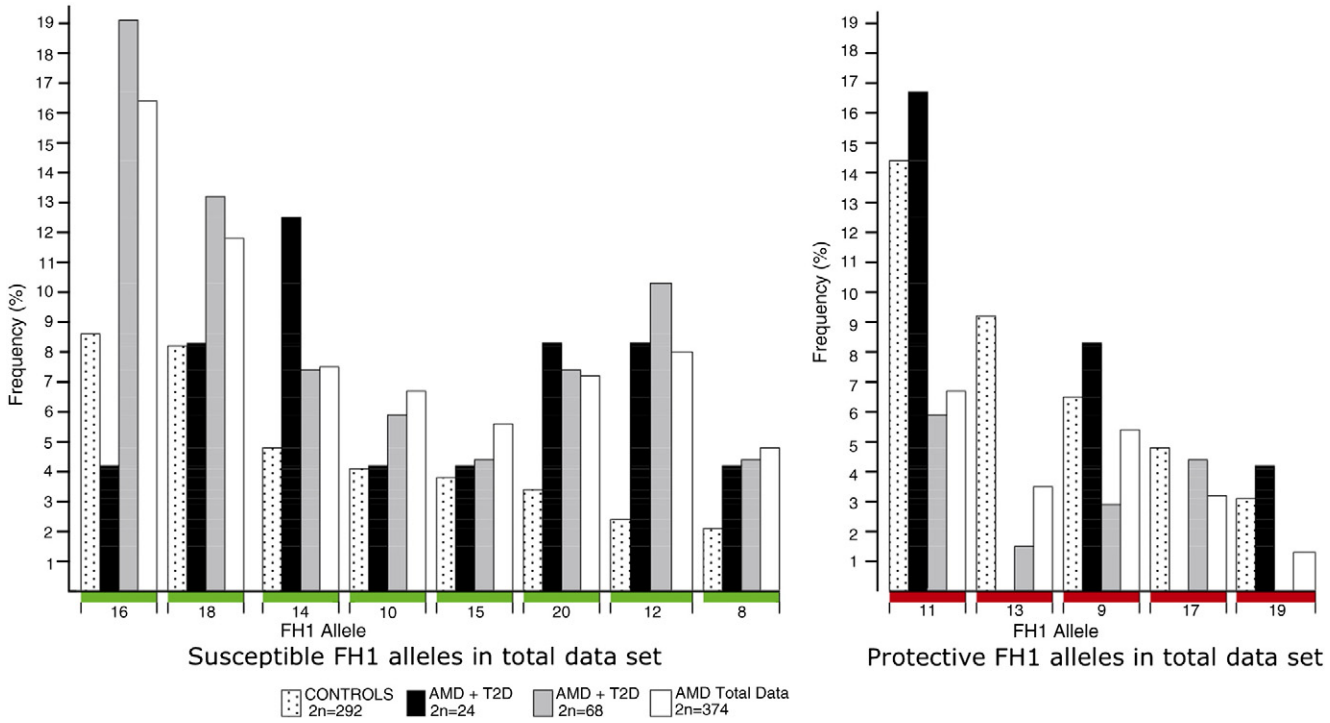


Fig. 6. Frequency distribution of FH1 alleles in AMD patients with and without diabetes shown in comparison with the total AMD patient and controls. T2D = Type 2 diabetes.

[9] J.B. Maller, J.A. Fagerness, R.C. Reynolds, B.M. Neale, M.J. Daly, J.M. Seddon, Variation in complement factor 3 is associated with risk of age-related macular degeneration, Nat. Genet. 39 (2007) 1200–1201.
[10] A.J. Richardson, F.M. Islam, R.H. Guymer, M. Cain, P.N. Baird, A tag-single nucleotide polymorphisms approach to the vascular endothelial growth factor-A gene in age-related macular degeneration, Mol. Vis. 13 (2007) 2148–2152.
[11] D.H. Anderson, M.J. Radeke, N.B. Gallo, E.A. Chapin, P.T. Johnson, C.R. Curletti, L.S. Hancox, J. Hu, J.N. Ebright, G. Malek, M.A. Hauser, C.B. Rickman, D. Bok, G.S. Hageman, L.V. Johnson, The pivotal role of the complement system in aging and age-related macular degeneration: hypothesis re-visited, Prog Retin Eye Res 29 (2010) 95–112.
[12] Y. Athanasiou, K. Voskarides, D.P. Gale, L. Damianou, C. Patsias, M. Zavros, P.H. Maxwell, H.T. Cook, P. Demosthenous, A. Hadjisavvas, K. Kyriacou, I. Zouvani, A. Pierides, C. Deltas, Familial C3 glomerulopathy associated with CFHR5 mutations: clinical characteristics of 91 patients in 16 pedigrees, Clin. J. Am. Soc. Nephrol. 6 (2011) 1436–1446.
[13] P.K. Kaiser, D.V. Do, Ranibizumab for the treatment of neovascular AMD, Int. J. Clin. Pract. 61 (2007) 501–509.
[14] K.L. Spencer, M.A. Hauser, L.M. Olson, N. Schnetz-Boutaud, W.K. Scott, S. Schmidt, P. Gallins, A. Agarwal, E.A. Postel, M.A. Pericak-Vance, J.L. Haines, Haplotypes spanning the complement factor H gene are protective against age-related macular degeneration, Invest. Ophthalmol. Vis. Sci. 48 (2007) 4277–4283.
[15] V. Romero, C.E. Larsen, J.S. Duke-Cohan, E.A. Fox, T. Romero, O.P. Clavijo, D.A. Fici, Z. Husain, I. Almeciga, D.R. Alford, Z.L. Awdeh, J. Zuniga, L. El-Dahdah, C.A. Alper, E.J. Yunis, Genetic fixity in the human major histocompatibility complex and block size diversity in the class I region including HLA-E, BMC Genet. 8 (2007) 14.
[16] R. Dawkins, C. Leelayuwat, S. Gaudieri, G. Tay, J. Hui, S. Cattley, P. Martinez, J. Kulski, Genomics of the major histocompatibility complex: haplotypes, duplication, retroviruses and disease, Immunol. Rev. 167 (1999) 275–304.
[17] S.K. Cattley, J.F. Williamson, G.K. Tay, O.P. Martinez, S. Gaudieri, R.L. Dawkins, Further characterization of MHC haplotypes demonstrates conservation telomeric of HLA-A: update of the 4AOH and 10 IHW cell panels, Eur. J. Immunogenet. 27 (2000) 397–426.
[18] S. Lester, C. McLure, J. Williamson, P. Bardy, M. Rischmueller, R.L. Dawkins, Epistasis between the MHC and the RCA alpha block in primary Sjogren syndrome, Ann. Rheum. Dis. 67 (2008) 849–854.

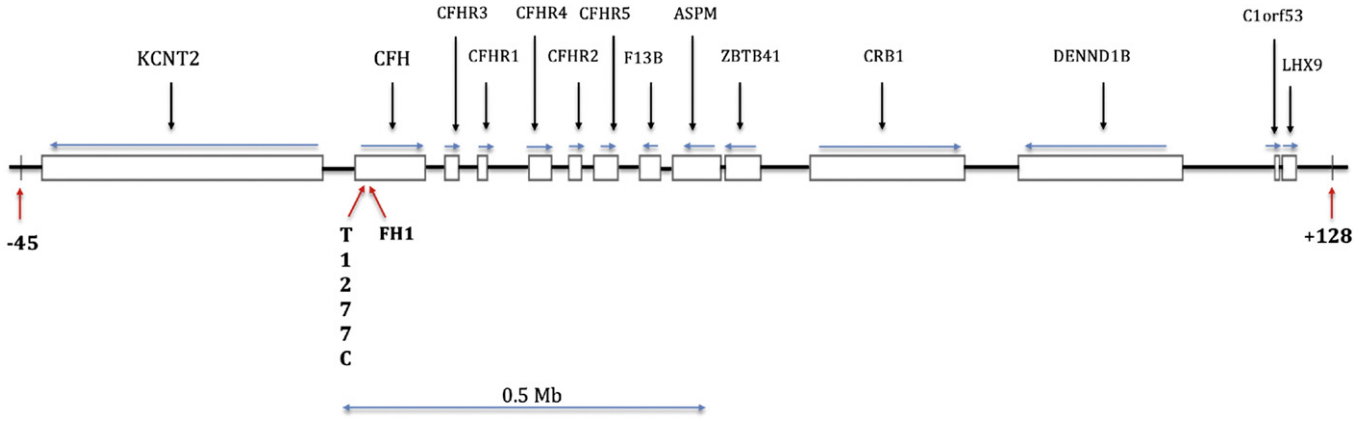


Fig. 7. Map of the location and distribution of genes and markers on chromosome 1q 31–32 RCA β (~196,114,000–197,900,000). The location of the four markers used to define RCA β haplotypes are shown under the line. Boxes represent genes with names shown above the line. KCNT2: potassium channel, subfamily T, member 2; CFH: complement factor H; CFHR1-5: complement factor H related genes; F13B: coagulation factor XIII B; ASPM: abnormal spindle (asp) homologue; ZBTB41: zink finger and BTB domain containing 41 homologue; CRB1: crumbs homologue 1; DENND1B: DENN/MADD domain containing 1B; C1orf53: chromosome 1 open reading frame 53; LHX9: LIM homeobox 9. Gene names and sequences were obtained from NCBI RefSeq Build 37.2 (http://www.ncbi.nlm.nih.gov/). Horizontal arrows indicate direction of transcription.

[19] E.J. Yunis, C.E. Larsen, M. Fernandez-Vina, Z.L. Awdeh, T. Romero, J.A. Hansen, C.A. Alper, Inheritable variable sizes of DNA stretches in the human MHC: conserved extended haplotypes and their fragments or blocks, Tissue Antigens 62 (2003) 1–20.
[20] D. Raum, Z. Awdeh, E.J. Yunis, C.A. Alper, K.H. Gabbay, Extended major histocompatibility complex haplotypes in type I diabetes mellitus, J. Clin. Invest. 74 (1984) 449–454.
[21] Z.L. Awdeh, D. Raum, E.J. Yunis, C.A. Alper, Extended HLA/complement allele haplotypes: evidence for T/t-like complex in man, Proc. Natl. Acad. Sci. U. S. A. 80 (1983) 259–263.
[22] J.F. Williamson, C.A. McLure, P.N. Baird, D. Male, J. Millman, B. Lawley, M.L. Ashdown, P.J. Keating, R.L. Dawkins, Novel sequence elements define ancestral haplotypes of the region encompassing complement factor H, Hum. Immunol. 69 (2008) 207–219.
[23] S. Rodriguez de Cordoba, D.M. Lublin, P. Rubinstein, J.P. Atkinson, Human genes for three complement components that regulate the activation of C3 are tightly linked, J. Exp. Med. 161 (1985) 1189–1195.
[24] C.A. Alper, C.E. Larsen, D.P. Dubey, Z.L. Awdeh, D.A. Fici, E.J. Yunis, The haplotype structure of the human major histocompatibility complex, Hum. Immunol. 67 (2006) 73–84.
[25] M.A. Degli-Esposti, A.L. Leaver, F.T. Christiansen, C.S. Witt, L.J. Abraham, R.L. Dawkins, Ancestral haplotypes: conserved population MHC haplotypes, Hum. Immunol. 34 (1992) 242–252.
[26] M.A. Degli-Esposti, L.J. Abraham, V. McCann, T. Spies, F.T. Christiansen, R.L. Dawkins, Ancestral haplotypes reveal the role of the central MHC in the immunogenetics of IDDM, Immunogenetics 36 (1992) 345–356.
[27] M.A. Degli-Esposti, A. Andreas, F.T. Christiansen, B. Schalke, E. Albert, R.L. Dawkins, An approach to the localization of the susceptibility genes for generalized myasthenia gravis by mapping recombinant ancestral haplotypes, Immunogenetics 35 (1992) 355–364.
[28] P.N. Baird, A. Richardson, A. Islam, L. Lim, R. Guymer, Analysis of the RDS/peripherin gene in age-related macular degeneration, Clin. Experiment. Ophthalmol. 35 (2007) 194–195.